Photo Uncrop

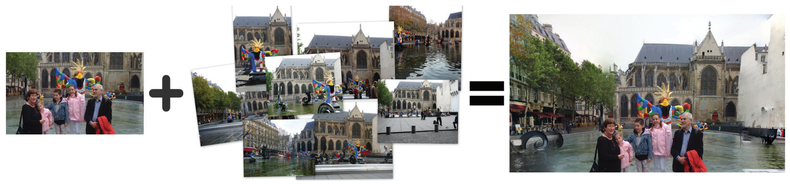
Qi Shan[1], Brian Curless[1], Yasutaka Furukawa[2], Carlos Hernandez[3], Steven M. Seitz[1,3]

[1]University of Washington [2]Washington University in St. Louis [3]Google Inc.

ECCV '14 Zürich, September 6-12th, 2014

Problem

Travel photos have limited field of view (FOV). We address the problem of extending the field of view of a photo – an operation we call ***uncrop***, by utilizing Internet photos from near-by viewpoints.



Challenges

(1) uncontrolled viewpoints and viewing directions – too much parallax for existing panorama stitchers;
(2) full 3D reconstruction not feasible for this application – rarely produces complete, high-res models with salient scene elements, e.g., people, cars, trees, and other transitory or hard-to-match objects;
(3) apperance variation in photos;
(4) ***dramatic*** boundary expansion range.

Assumptions

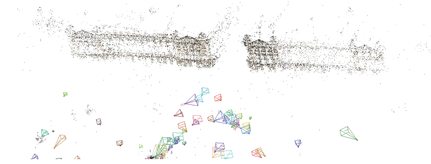
(1) user input photos are taken at sites where a sufficient number of Internet photos are available;
(2) for the application of extending spatial context for a photo, small artifacts are tolerable.

Workflow

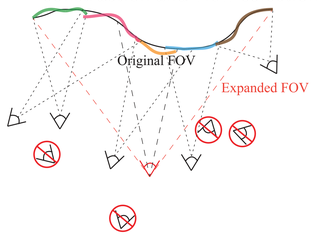
(0) Input data

User input Internet photos

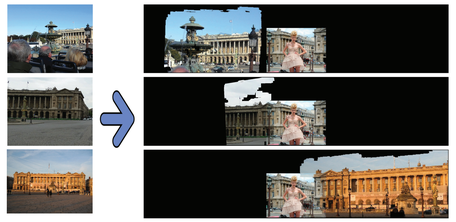
(1) Structure from Motion + Multi-view Stereo



(2) Dense per-view depth map construction and view selection



(3) Image warping



(4) Construct an MRF-based composite (see next panel)

(5) Poisson blending

Input image MRF composite Final composite after Poisson blending

MRF formulation

p is a pixel, l is the label map

$$E(l) = \sum_p E_{\text{unary}}(p, l(p)) + \sum_{\{p,q\}\in\mathcal{N}(p,q)} E_{\text{binary}}(p, l(p), q, l(q)) + E_{\text{label}}(l)$$

$E_{\text{label}}(l)$: penalize large number of labels

$$E_{\text{unary}}(p,l) = E_{\text{geometry}}(p,l) + \alpha_1 E_{\text{appearance}}(p,l) + \alpha_2 E_{\text{contrast}}(p,l) + \alpha_3 E_{\text{reference}}(p,l)$$

$E_{\text{appearance}}$: favor images with similar appearance to the reference
E_{contrast}: favor high contrast pixels
$E_{\text{reference}}$: keep reference image pixels

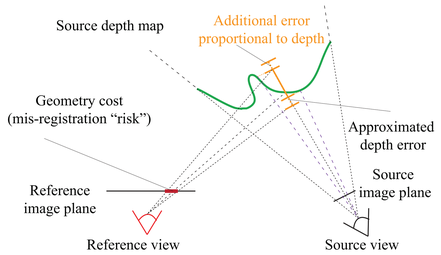
$$E_{\text{binary}} = E_{\text{edge}} + \beta E_{\text{compatibility}}$$

E_{edge}: favor cuts at less noticeable locations

Novel terms:

E_{geometry}

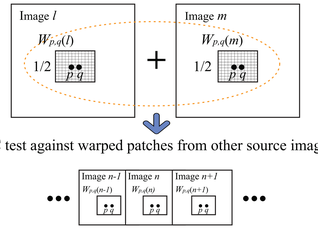
Measures uncertainty of reprojected pixel location, i.e., potential misregistration. We use depth variation around a source pixel to model depth uncertainty before reprojection.

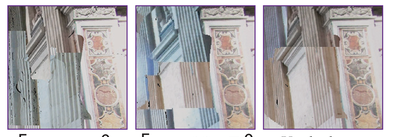


$E_{\text{compatibility}}$

To encourage regions in the target image to resemble structures in the source images.

$$E_{\text{compatibility}}(p,l,q,m) = 1 - \max_n \text{NCC}\left[\frac{1}{2}\left(W_{p,q}(l) + W_{p,q}(m)\right), W_{p,q}(n)\right]$$





$E_{\text{geometry}} = 0$ $E_{\text{compatibility}} = 0$ Use both terms

Results

Evaluated on 10 datasets in the city of Rome and Paris, number of images ranges from 262 to 2397.

User input

Photoshop CS6 PhotoMerge with manual color blending

Median composite (the composite step in [Zhang et al. 2014])

[Nomura et al. 2007]

Our *photo uncrop* result

Zhang, C., Gao, J., Wang, O., Georgel, P., Yang, R., Davis, J., Frahm, J.M., Pollefeys, M.: Personal photo enhancement using internet photo collections. TVCG (2014).
Nomura, Y., Zhang, L., Nayar, S.: Scene collages and flexible camera arrays. In: Eurographics Symposium on Rendering (2007).